

# QUALITATIVE NOISE-DISTURBANCE RELATION FOR QUANTUM MEASUREMENTS

TEIKO HEINOSAARI<sup>†</sup> AND TAKAYUKI MIYADERA<sup>‡</sup>

ABSTRACT. The inherent connection between noise and disturbance is one of the most fundamental features of quantum measurements. In the two well-known extreme cases a measurement either makes no disturbance but then has to be totally noisy or is as accurate as possible but then has to disturb so much that all subsequent measurements become redundant. Most of the measurements are, however, something between these two extremes. We derive a structural relation between observables and channels that properly explains the trade-off between noise and disturbance.

## 1. INTRODUCTION

The inherent connection between noise and disturbance is one of the most fundamental features of quantum measurements. On the one hand, a measurement cannot give any information without disturbing the object system. On the other hand, a noisier (less informative) measurement can be implemented with less disturbance than a sharper measurement. Roughly speaking, more noise means that measurement outcome distributions become broader, while disturbance is reflected in the measurement outcome statistics of subsequent measurements. In the most extreme case, the disturbance inherent in a measurement makes all subsequent measurements irrelevant.

Various trade-off inequalities between noise (or information) and disturbance are known, all depending on different quantification of these notions, see e.g. [1, 2, 3, 4, 5, 6]. All these trade-off inequalities are revealing different aspects of the interplay between noise and disturbance in quantum measurements. In this work we present a relation between noise and disturbance which is qualitative in nature and not based on any specific quantifications of noise and disturbance. Our result is a structural connection between observables and channels. More precisely, we show that a certain partial order in the set of equivalence classes of quantum observables (positive operator valued measures) corresponds to an inclusion of the related subsets of quantum channels (trace preserving completely positive maps). As we will explain,

this correspondence has a clear interpretation as a noise-disturbance relationship since it shows how the possible state transformations are limited to more noisy ones if the measurement is required to be more accurate. Due to its simplicity and generality, we believe that our qualitative noise-disturbance relation can be seen as a common origin of many quantitative noise-disturbance inequalities.

To give a preliminary idea on the coming developments, we recall two well-known special situations. (See e.g. [7, 8] for general results that cover these cases.) First, let us consider a measurement in an orthonormal basis  $\{\varphi_j\}_{j=1}^d$ . If  $\varrho$  is an input state, then the measurement outcome probabilities are  $\langle \varphi_j | \varrho \varphi_j \rangle$ . The output state is a mixture  $\sum_j \langle \varphi_j | \varrho \varphi_j \rangle \xi_j$ , where  $\xi_1, \xi_2, \dots$  are states that depend on the measurement device but not on the input state. Hence, a measurement in an orthonormal basis is sharp but disturbs a lot. A completely different kind of measurement is such that we do nothing on the input state but we just throw a dice to produce measurement outcome probabilities. This measurement has maximum amount of noise, but it can be implemented without disturbing the input state at all.

Most of measurements belong to the intermediate area between the two previously described extreme cases. Namely, they contain some additional noise and can be measured in a way that implies some disturbance. More noise should mean less disturbance, and vice versa. It is exactly this kind of intuitive trade-off that we will turn into an exact theorem.

In the rest of the paper  $\mathcal{H}$  is a fixed Hilbert space related to the input system. The dimension of  $\mathcal{H}$  can be either finite or countably infinite. We denote by  $\mathcal{L}(\mathcal{H})$  the set of all bounded operators on  $\mathcal{H}$ . A quantum measurement produces measurement outcomes and conditional output states. The mapping from input states into measurement outcome statistics is called an observable, while the mapping from input states into unconditional output states (i.e. average over conditional output states) is called a channel [9]. We will briefly recall some of the basic properties of observables and channels before proving our main result, Theorem 2.

## 2. ORDER STRUCTURE OF OBSERVABLES

A quantum observable with finite or countably infinite number of outcomes is described by a mapping  $x \mapsto A(x)$  such that each  $A(x) \in \mathcal{L}(\mathcal{H})$  is a positive operator (i.e.  $\langle \psi | A(x) \psi \rangle \geq 0$  for all  $\psi \in \mathcal{H}$ ) and  $\sum_x A(x) = \mathbb{1}$ , where  $\mathbb{1}$  is the identity operator on  $\mathcal{H}$ . The labeling of measurement outcomes is not important for the questions that we will

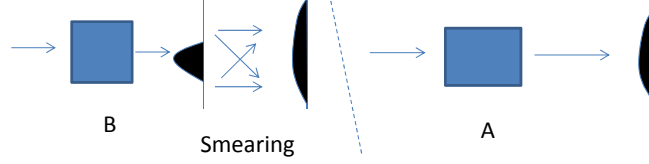


FIGURE 1. If  $A \preceq B$ , then a measurement of  $A$  can be simulated by a measurement of  $B$  and a classical channel  $M$  applied to the measurement outcome distribution.

investigate, hence we assume that the outcome set of all our observables is  $\mathbb{N} = \{1, 2, \dots\}$ . We denote by  $\mathfrak{O}$  the set of all observables on  $\mathcal{H}$ . Let us remark that it is possible that  $A(x) = 0$  for some outcomes  $x$ , hence e.g. observables with only finite number of outcomes are included in  $\mathfrak{O}$  by adding zero operators. For each observable  $A$ , we denote by  $\Omega_A \subseteq \mathbb{N}$  the set of all outcomes  $x$  with  $A(x) \neq 0$ .

By a *stochastic matrix* we mean a real matrix  $[M_{xy}]$ ,  $x, y \in \mathbb{N}$  such that  $M_{xy} \geq 0$  and  $\sum_x M_{xy} = 1$ . Given two observables  $A$  and  $B$ , we denote  $A \preceq B$  if there exists a stochastic matrix  $M$  such that

$$A(x) = \sum_y M_{xy} B(y). \quad (1)$$

The relation  $\preceq$  is a preordering in  $\mathfrak{O}$ , i.e.,  $A \preceq A$  for every observable  $A$ , and if  $A \preceq B$  and  $B \preceq C$ , then  $A \preceq C$ . This preordering structure has been called with different names in the literature; non-ideality [10], smearing [11], post-processing [12]. The physical meaning of the relation is that if  $A \preceq B$ , then (in the level of measurement outcome statistics) a measurement of  $A$  can be simulated by a measurement of  $B$  and a classical channel applied to the measurement outcome distribution; see Fig. 1. In this sense,  $B$  is superior to  $A$ . The physical mechanism of the additional noise of  $A$  compared to  $B$  is typically related to a weaker measurement coupling or impurities in the ancilla state. We refer to [11] for some realistic examples.

Let us notice that it is possible to have  $A \preceq B$  and  $B \preceq A$  even if  $A \neq B$  [13]. For this reason, it is often appropriate to study equivalence classes of observables rather than single observables. We denote  $A \simeq B$  if and only if both  $A \preceq B$  and  $B \preceq A$  hold. Then  $\simeq$  is an equivalence relation and the equivalence class of  $A$  is denoted by  $[A]$ . Physically speaking, the equivalence class  $[A]$  contains all observables  $B$  that are like  $A$  in all relevant ways but may differ by the ordering of measurement outcomes or some other irrelevant detail. We introduce the set of equivalence classes  $\mathfrak{O}^\sim := \mathfrak{O} / \simeq$  and the preorder  $\preceq$  then

induces a partial order  $\preceq$  on  $\mathfrak{D}^\sim$  by  $[A] \preceq [B]$  if and only if  $A \preceq B$ . (We use the same symbol  $\preceq$  for these two different relations, but this should not cause a confusion.) It is easy to see that in the partially ordered set  $\mathfrak{D}^\sim$ , there exists the least element but there is no greatest element. Namely, an observable  $C$  defined by  $C(1) = \mathbb{1}$ ,  $C(j) = 0$  for  $j \neq 1$  is a representative of the least element since for every  $A \in \mathfrak{D}$ , the equality  $\mathbb{1} = \sum_x A(x)$  holds. The equivalence class  $[C]$  consists of all 'coin tossing observables', i.e.,

$$[C] = \{C_p | C_p(x) = p(x)\mathbb{1}, 0 \leq p(x) \leq 1, \sum_x p(x) = 1\}.$$

The measurement outcome of an observable  $C_p$  is determined by a fixed probability distribution  $p$  and does not depend on the input state at all.

To see that there is no greatest element in  $\mathfrak{D}^\sim$ , suppose on the contrary that  $B$  is such. Let  $\{\varphi_x\}$  be an orthonormal basis and define an observable  $A$  by  $A(x) = |\varphi_x\rangle\langle\varphi_x|$ . Then the condition  $|\varphi_x\rangle\langle\varphi_x| = \sum_y M_{xy} B(y)$  implies that every  $B(y)$  is proportional to some  $|\varphi_x\rangle\langle\varphi_x|$ . But since this should hold for arbitrary orthonormal basis  $\{\varphi_x\}$ , we must have  $B(y) = 0$ . This contradicts the fact that  $\sum_y B(y) = \mathbb{1}$ .

### 3. ORDER STRUCTURE OF CHANNELS

A measurement process yields a probability distribution of measurement outcomes, but it also causes a change of the input state. This state transformation is described by a quantum channel. In the Schrödinger picture a channel is a completely positive map that maps an input state to an output state. We allow the output state to belong to a different operator space  $\mathcal{L}(\mathcal{K})$  than the input state. For instance, a mapping  $\varrho \mapsto \varrho \otimes \xi$ , where  $\xi \in \mathcal{L}(\mathcal{K})$  is a fixed state, is a valid channel. This particular channel adds an ancilla system in a state  $\xi$  to the original system.

For the purposes of this paper, it is more convenient to use the Heisenberg picture description for channels. In the Heisenberg picture a channel is defined as a normal completely positive map  $\Lambda : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$  satisfying  $\Lambda(\mathbb{1}_{\mathcal{K}}) = \mathbb{1}_{\mathcal{H}}$ , where  $\mathcal{K}$  is the output Hilbert space. The Schrödinger picture description  $\Lambda^S$  of a channel  $\Lambda$  can be obtained from the relation

$$\text{tr} [\Lambda^S(\varrho)C] = \text{tr} [\varrho\Lambda(C)], \quad (2)$$

true for all states  $\varrho \in \mathcal{L}(\mathcal{H})$  and operators  $C \in \mathcal{L}(\mathcal{K})$ .

We denote by  $\mathfrak{C}$  the set of all channels from an arbitrary output space  $\mathcal{L}(\mathcal{K})$  to the fixed input space  $\mathcal{L}(\mathcal{H})$ . For two channels  $\Lambda_1, \Lambda_2 \in \mathfrak{C}$ , we

denote  $\Lambda_1 \lesssim \Lambda_2$  if there exists a channel  $\mathcal{E}$  such that  $\Lambda_1 = \Lambda_2 \circ \mathcal{E}$ . This is analogous relation than the one defined for observables, and the physical meaning of  $\Lambda_1 \lesssim \Lambda_2$  is that  $\Lambda_1$  can be simulated by using  $\Lambda_2$  and  $\mathcal{E}$  sequentially. It is easy to see that this relation is a preorder but not a partial order.

As in the case of observables, it is often convenient to work on the level of equivalence classes of channels. If  $\Lambda_1 \lesssim \Lambda_2$  and  $\Lambda_2 \lesssim \Lambda_1$  hold, then we denote  $\Lambda_1 \sim \Lambda_2$ . The relation  $\sim$  is an equivalence relation, which allows us to introduce the set of equivalence classes  $\mathfrak{C}^\sim := \mathfrak{C} / \sim$ . The equivalence class of a channel  $\Lambda$  is denoted by  $[\Lambda] \in \mathfrak{C}^\sim$ , and a natural partial order  $\lesssim$  is introduced by  $[\Lambda_1] \lesssim [\Lambda_2]$  if and only if  $\Lambda_1 \lesssim \Lambda_2$ .

In the partially order set  $\mathfrak{C}^\sim$ , there exists the greatest element and the least element. Namely, for a state  $\varrho \in \mathcal{L}(\mathcal{H})$ , we define

$$\Lambda_\varrho : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H}), \quad \Lambda_\varrho(C) = \text{tr}[\varrho C] \mathbb{1}_{\mathcal{H}}. \quad (3)$$

Then for any  $\Lambda : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$ , the equation  $\Lambda_\varrho = \Lambda \circ \Lambda'_\varrho$  holds, where  $\Lambda'_\varrho : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$  is defined as  $\Lambda'_\varrho(C) = \text{tr}[\varrho C] \mathbb{1}_{\mathcal{K}}$ . Thus  $[\Lambda_\varrho]$  is the least element in  $\mathfrak{C}^\sim$ . On the other hand, the identity channel  $id : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$  defined by  $id(C) = C$  for all  $C \in \mathcal{L}(\mathcal{H})$  belongs to the greatest equivalence class since any channel  $\Lambda$  satisfies  $\Lambda = id \circ \Lambda$ .

#### 4. COMPATIBLE OBSERVABLES AND CHANNELS

A unifying description of the measurement outcome statistics and the state change under a measurement process is given by the notion of an instrument [14]. In the Schrödinger picture an instrument is a mapping  $(x, \varrho) \mapsto \mathcal{I}_x^S(\varrho)$  such that  $\text{tr}[\mathcal{I}_x^S(\varrho)]$  is the probability of obtaining an outcome  $x$  and the operator  $\tilde{\varrho}_x = \mathcal{I}_x^S(\varrho) / \text{tr}[\mathcal{I}_x^S(\varrho)]$  is the conditional output state under the condition that a measurement outcome  $x$  is obtained. The unconditional output state is thus given by  $\tilde{\varrho} \equiv \sum_x \mathcal{I}_x^S(\varrho)$ . The map  $\varrho \mapsto \tilde{\varrho}$  is a channel in the Schrödinger picture. We recall that every instrument has a measurement model consisting of an ancillary system and its initial state, a measurement interaction and a pointer observable on the ancillary system [15]. As in the case of channels, the Heisenberg picture for instruments is convenient for our purposes. An instrument in the Heisenberg picture is defined by a family of normal completely positive maps  $\mathcal{I}_x : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$  whose sum  $\sum_x \mathcal{I}_x$  is a channel.

We are interested in what pairs of observables and channels can belong to a same measurement process. Therefore, the following concept is useful.

**Definition 1.** Let  $A$  be an observable on  $\mathcal{L}(\mathcal{H})$ . A channel  $\Lambda : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$  is  $A$ -channel if there exists an instrument  $\mathcal{I}$  such that

$$\mathcal{I}_x(\mathbb{1}_{\mathcal{K}}) = A(x), \quad \sum_x \mathcal{I}_x(C) = \Lambda(C).$$

We denote by  $\mathfrak{C}_A$  the set of all  $A$ -channels.

In other words,  $\Lambda$  is  $A$ -channel if  $\Lambda$  and  $A$  are parts of a single instrument  $\mathcal{I}$ . Following [16], we call such  $\Lambda$  and  $A$  *compatible*.

Let  $A$  be an observable on  $\mathcal{L}(\mathcal{H})$ . If  $\Lambda \in \mathfrak{C}$  is an  $A$ -channel, any  $\Lambda' \in \mathfrak{C}$  satisfying  $\Lambda' \lesssim \Lambda$  is also an  $A$ -channel. Namely, suppose there exists an instrument  $\mathcal{I}$  such that  $\Lambda = \sum_x \mathcal{I}_x$  and  $\mathcal{I}_x(\mathbb{1}) = A(x)$ . If  $\Lambda' = \Lambda \circ \mathcal{E}$  for some channel  $\mathcal{E}$ , then we have  $\Lambda' = \sum_x \mathcal{I}_x \circ \mathcal{E}$  and  $(\mathcal{I}_x \circ \mathcal{E})(\mathbb{1}) = A(x)$ . Consequently, if  $\Lambda$  is an  $A$ -channel, any  $\Lambda' \in [\Lambda]$  is also an  $A$ -channel. Thus, a subset  $\mathfrak{C}_A^\sim$  of  $\mathfrak{C}^\sim$  is naturally introduced as  $\mathfrak{C}_A^\sim = \{[\Lambda] \mid \Lambda \text{ is an } A\text{-channel}\}$ . It is easy to see that the partially ordered set  $\mathfrak{C}_A^\sim$  contains the least element. Namely,  $\mathfrak{C}_A^\sim$  contains the least element of  $\mathfrak{C}^\sim$ , the equivalence class  $[\Lambda_\varrho]$ , introduced in (3). The fact that  $\Lambda_\varrho$  belongs to  $\mathfrak{C}_A$  for any observable  $A$  relates to the possibility of performing a destructive measurement; we can always measure  $A$ , destroy the system and prepare a state  $\varrho$ .

A less obvious and more interesting fact is that the partially ordered set  $\mathfrak{C}_A^\sim$  contains the greatest element. To construct a channel belonging to the greatest element of  $\mathfrak{C}_A^\sim$ , let  $(\mathcal{K}, \hat{A}, K)$  be a Naimark extension of  $A$  [17]. That is,  $\mathcal{K}$  is a Hilbert space,  $K : \mathcal{H} \rightarrow \mathcal{K}$  is an isometry, and  $\hat{A}$  is a projection-valued measure (PVM) on  $\mathcal{K}$  satisfying  $K^* \hat{A}(x) K = A(x)$  for all  $x \in \mathbb{N}$ . We define a channel  $\Lambda_A : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$  by

$$\Lambda_A(C) = \sum_x K^* \hat{A}(x) C \hat{A}(x) K. \quad (4)$$

To see that  $\Lambda_A$  is an  $A$ -channel, we define an instrument  $\mathcal{I}$  by

$$\mathcal{I}_x(C) = K^* \hat{A}(x) C \hat{A}(x) K. \quad (5)$$

Then  $\sum_x \mathcal{I}_x = \Lambda_A$  and  $\mathcal{I}_x(\mathbb{1}) = K^* \hat{A}(x) K = A(x)$ . Although the construction of  $\Lambda_A$  relies on the choice of the Naimark extension  $(\mathcal{K}, \hat{A}, K)$ , the following arguments do not depend on this choice. From now on, we will always assume that a Naimark extension  $(\mathcal{K}, \hat{A}, K)$  has been fixed for each observable  $A$ , hence also  $\Lambda_A$  is defined for each  $A$ .

**Theorem 1.** Let  $A$  be an observable. The set  $\mathfrak{C}_A$  of all  $A$ -channels consists of all channels that are below  $\Lambda_A$ , i.e.,

$$\mathfrak{C}_A = \{\Lambda \in \mathfrak{C} \mid \Lambda \lesssim \Lambda_A\}. \quad (6)$$

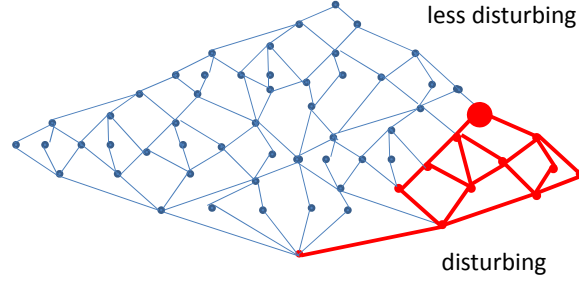


FIGURE 2. The set  $\mathfrak{C}_A^\sim$  (red) consists of all elements that are below a single element  $[\Lambda_A]$  (big dot).

Thus,  $\mathfrak{C}_A^\sim$  has the greatest element  $[\Lambda_A]$  and

$$\mathfrak{C}_A^\sim = \{[\Lambda] \in \mathfrak{C}^\sim \mid [\Lambda] \lesssim [\Lambda_A]\}. \quad (7)$$

From the mathematical point of view, the set  $\mathfrak{C}_A^\sim$  generated by a single element  $[\Lambda_A]$  is called a principal ideal, which is the minimal ideal containing  $[\Lambda_A]$ . The result of Theorem 1 is illustrated in Fig. 2.

**Proof of Theorem 1.** We have already seen that  $\mathfrak{C}_A \supseteq \{\Lambda \in \mathfrak{C} \mid \Lambda \lesssim \Lambda_A\}$ , hence we need to show that the inclusion holds in the other direction as well.

Let  $\Lambda : \mathcal{L}(\mathcal{K}') \rightarrow \mathcal{L}(\mathcal{H})$  be an  $A$ -channel. To prove that  $\Lambda \lesssim \Lambda_A$ , we first fix a minimal Stinespring dilation  $(\mathcal{K}'', V)$  of  $\Lambda$ . Thus,  $\mathcal{K}''$  is a Hilbert space,  $V : \mathcal{H} \rightarrow \mathcal{K}' \otimes \mathcal{K}''$  is an isometry satisfying  $\Lambda(C) = V^*(C \otimes \mathbb{1})V$  and the set  $(\mathcal{L}(\mathcal{K}') \otimes \mathbb{1})V\mathcal{H}$  is dense in  $\mathcal{K}' \otimes \mathcal{K}''$ . Since  $\Lambda$  is an  $A$ -channel, we can apply the Radon-Nikodym theorem of CP-maps [18, 19] to conclude that there exists a unique observable  $R$  on  $\mathcal{L}(\mathcal{K}'')$  satisfying

$$A(x) = V^*(\mathbb{1} \otimes R(x))V$$

for all  $x \in \mathbb{N}$ . For each  $x \in \Omega_A$ , we define an operator  $c_x : \mathcal{H} \rightarrow \mathcal{K}' \otimes \mathcal{K}''$  by

$$c_x := (\mathbb{1} \otimes R(x)^{1/2})V.$$

Then for any  $C \in \mathcal{L}(\mathcal{K}')$ , we have

$$\Lambda(C) = \sum_x c_x^*(C \otimes \mathbb{1})c_x. \quad (8)$$

Since  $c_x$  satisfies  $c_x^*c_x = A(x)$ , by the polar decomposition theorem there exists an isometry  $W_x : \mathcal{H} \rightarrow \mathcal{K}' \otimes \mathcal{K}''$  satisfying

$$c_x = W_x \sqrt{A(x)}, \quad (9)$$

and therefore

$$\Lambda(C) = \sum_x \sqrt{A(x)} W_x^* (C \otimes \mathbb{1}) W_x \sqrt{A(x)}. \quad (10)$$

We note that if  $\dim \mathcal{H} = \infty$ , then the polar decomposition theorem states that  $W_x$  is a partial isometry (and not necessarily isometry). However, in our setting it is possible to extend the partial isometry to an isometric operator. This additional argument is given in Appendix.

Let  $(\mathcal{K}, \hat{A}, K)$  be the Naimark dilation of  $A$ . The relationship  $K^* \hat{A}(x) K = A(x)$  implies that there exists an isometry  $J_x : \mathcal{H} \rightarrow \mathcal{K}$  satisfying

$$\hat{A}(x) K = J_x \sqrt{A(x)}. \quad (11)$$

Again, the argument why  $J_x$  is an isometry and not just a partial isometry is given in Appendix. Inserting (11) into (10) gives

$$\Lambda(C) = \sum_x K^* \hat{A}(x) J_x W_x^* (C \otimes \mathbb{1}) W_x J_x^* \hat{A}(x) K.$$

Finally, fix an arbitrary state  $\rho$  on  $\mathcal{K}'$ . We define

$$\begin{aligned} \mathcal{E}(C) &:= \sum_x \hat{A}(x) J_x W_x^* (C \otimes \mathbb{1}_{\mathcal{K}'}) W_x J_x^* \hat{A}(x) \\ &\quad + \text{tr}[\rho C] \left( \mathbb{1} - \sum_x \hat{A}(x) J_x J_x^* \hat{A}(x) \right). \end{aligned}$$

Then  $\mathcal{E}$  is a channel and

$$\begin{aligned} \Lambda_A \circ \mathcal{E}(C) &= \Lambda(C) + \\ &+ \text{tr}[\rho C] \left( \sum_x K^* \hat{A}(x) K - \sum_x K^* \hat{A}(x) J_x J_x^* \hat{A}(x) K \right) \\ &= \Lambda(C) + \text{tr}[\rho C] \left( \mathbb{1} - \sum_x \sqrt{A(x)} \sqrt{A(x)} \right) = \Lambda(C). \end{aligned}$$

Thus we obtain  $\Lambda = \Lambda_A \circ \mathcal{E}$ , implying that  $\Lambda \lesssim \Lambda_A$ . ■

Suppose that  $A$  and  $B$  are two observables satisfying  $\mathfrak{C}_B \subseteq \mathfrak{C}_A$ . Even without any quantification of noise, we can obviously say that it is possible to measure  $A$  with less or equal disturbance than generated in any measurement of  $B$ . In other words, the unavoidable disturbance related to  $A$  is smaller than equal to the unavoidable disturbance related to  $B$ . This qualitative description of disturbance will be the basis of the forthcoming noise - disturbance relation.



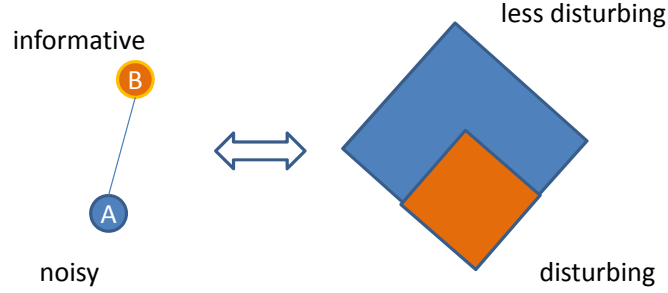


FIGURE 3. Illustration of Theorem 2:  $A \preceq B$  (left) if and only if  $\mathfrak{C}_A \supseteq \mathfrak{C}_B$  (right).

### 5. NOISE – DISTURBANCE RELATION

The following preliminary observation is easily extracted from our earlier discussion and Theorem 1.

**Lemma 1.** *Let  $A$  and  $B$  be two observables. Then  $\mathfrak{C}_B \subseteq \mathfrak{C}_A$  if and only if  $\Lambda_B \in \mathfrak{C}_A$ .*

We are now ready to proceed to our main result.

**Theorem 2.** *Let  $A$  and  $B$  be two observables. Then  $A \preceq B$  if and only if  $\mathfrak{C}_B \subseteq \mathfrak{C}_A$ .*

This result is illustrated in Fig. 3.

**Proof of Theorem 2.** The *only if*-part: Suppose that  $A \preceq B$ , hence there exists a stochastic matrix  $M$  such that  $A(x) = \sum_y M_{xy} B(y)$ . Let  $\Lambda : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$  be a  $B$ -channel, meaning that there exists an instrument  $\mathcal{I}$  such that

$$\mathcal{I}_y(\mathbb{1}_{\mathcal{K}}) = B(y), \quad \sum_y \mathcal{I}_y(C) = \Lambda(C).$$

We define an instrument  $\mathcal{I}'$  by formula  $\mathcal{I}'_x := \sum_y M_{xy} \mathcal{I}_y$ . Then it is easy to see  $\sum_x \mathcal{I}'_x = \Lambda$  and  $\mathcal{I}'_x(\mathbb{1}_{\mathcal{K}}) = A(x)$ . Therefore,  $\Lambda$  is an  $A$ -channel. Since  $\Lambda$  was an arbitrary  $B$ -channel, we conclude that  $\mathfrak{C}_B \subseteq \mathfrak{C}_A$ .

The *if*-part: By Lemma 1 we have  $\Lambda_B \in \mathfrak{C}_A$ . A Stinespring representation of  $\Lambda_B$  is given by an isometry  $V : \mathcal{H} \rightarrow \mathcal{K} \otimes \mathcal{K}'$ ,

$$V\psi = \sum_{x \in \Omega_B} \hat{B}(x) K\psi \otimes e_x,$$

where  $\mathcal{K}'$  is a Hilbert space with the dimension equal to the cardinality of  $\Omega_B$  and  $\{e_x\}$  is an orthonormal basis of  $\mathcal{K}'$ . If the channel  $\Lambda_B$  is compatible with the observable  $A$ , then from the Radon-Nikodym theorem

of CP-maps [18, 19] follows that there exists an observable  $\mathbf{Y}$  acting on  $\mathcal{K}'$  such that

$$\mathbf{A}(y) = V^*(\mathbb{1} \otimes \mathbf{Y}(y))V$$

for all  $y \in \mathbb{N}$ . (In case the Stinespring representation is not minimal, the uniqueness of  $\mathbf{Y}$  drops.) Thus we obtain for any  $\psi \in \mathcal{H}$ ,

$$\begin{aligned} \langle \psi | \mathbf{A}(y) \psi \rangle &= \sum_x \sum_{x'} (\hat{B}(x)K\psi, \hat{B}(x')K\psi)(e_x, \mathbf{Y}(y)e_{x'}) \\ &= \sum_x (\psi, K^* \hat{B}(x)K\psi)(e_x, \mathbf{Y}(y)e_x) \\ &= (\psi, \sum_x \mathbf{B}(x)(e_x, \mathbf{Y}(y)e_x)\psi), \end{aligned}$$

where we used  $\hat{B}(x)\hat{B}(x') = \delta_{xx'}\hat{B}(x)$ . As  $M_{yx} := (e_x, \mathbf{Y}(y)e_x)$  is a stochastic matrix, we conclude that  $\mathbf{A} \preceq \mathbf{B}$ . ■

## 6. EXAMPLE: BINARY QUBIT MEASUREMENTS

The simplest kind of measurements are binary (i.e. two-outcome) measurements on a qubit system. For each vector  $\vec{v} \in \mathbb{R}^3$  with  $\|\vec{v}\| \leq 1$ , we define a binary qubit observable  $\mathbf{A}^{\vec{v}}$  by

$$\mathbf{A}^{\vec{v}}(\pm 1) = \frac{1}{2}(\mathbb{1} \pm \vec{v} \cdot \vec{\sigma}).$$

It is easy to see that  $\mathbf{A}^{\vec{w}} \preceq \mathbf{A}^{\vec{v}}$  if and only if  $\vec{w}$  and  $\vec{v}$  are parallel vectors and  $\|\vec{w}\| \leq \|\vec{v}\|$ . To demonstrate how this order structure of observables is reflected in the measurement disturbance, let us consider the Lüders measurements for the above type of qubit observables. The Lüders instrument related to  $\mathbf{A}^{\vec{v}}$  is defined as

$$\mathcal{I}_x^{\vec{v}}(C) = \sqrt{\mathbf{A}^{\vec{v}}(x)}C\sqrt{\mathbf{A}^{\vec{v}}(x)}, \quad x = \pm 1.$$

The corresponding channel is

$$\Lambda^{\vec{v}} = \mathcal{I}_1^{\vec{v}} + \mathcal{I}_{-1}^{\vec{v}} = \lambda \text{id} + (1 - \lambda) \mathcal{V}, \quad (12)$$

where

$$\mathcal{V}(C) = 1/\|\vec{v}\|^2 \vec{v} \cdot \vec{\sigma} C \vec{v} \cdot \vec{\sigma}, \quad \lambda = \frac{1 + \sqrt{1 - \|\vec{v}\|^2}}{2}. \quad (13)$$

Let us notice that the unitary channel  $\mathcal{V}$  depends on the direction of  $\vec{v}$  but not on its norm, while the weight  $\lambda$  depends on the norm of  $\vec{v}$  but not on its direction. Applying Theorem 2 for two observables  $\mathbf{A}^{\vec{v}}$  and  $\mathbf{A}^{\vec{w}}$  with parallel vectors  $\vec{v}$  and  $\vec{w}$ , we conclude that for two parameters

$\lambda, \mu \in [\frac{1}{2}, 1]$  and a unitary channel  $\mathcal{V}$  defined in (13), there exists a channel  $\mathcal{E}$  such that

$$(\lambda \text{ id} + (1 - \lambda) \mathcal{V}) \circ \mathcal{E} = (\mu \text{ id} + (1 - \mu) \mathcal{V}). \quad (14)$$

if and only if  $\lambda \geq \mu$ . This is in line what we would expect; the more sharper is the measurement, the smaller must be the weight of the identity channel. In this example, it is not too difficult to find the concrete form of a channel  $\mathcal{E}$  satisfying (14). Namely, for all  $\lambda, \lambda' \in [\frac{1}{2}, 1]$ , we obtain

$$\begin{aligned} & (\lambda \text{ id} + (1 - \lambda) \mathcal{V}) \circ (\lambda' \text{ id} + (1 - \lambda') \mathcal{V}) = \\ & ((1 - \lambda - \lambda' + 2\lambda\lambda') \text{ id} + (\lambda + \lambda' - 2\lambda\lambda') \mathcal{V}). \end{aligned} \quad (15)$$

Hence, for every  $\mu < \lambda$  we can choose  $\lambda' = (\mu + \lambda - 1)/(2\lambda - 1)$  and then (15) leads to (14).

## 7. SUMMARY

Classical and quantum post-processings yield physically meaningful relations in the sets of observables and channels, respectively. When lifted to the sets of equivalence classes, these relations become partial orderings. The partial orderings can be seen as abstract and general ways to describe noise and disturbance. We have proved that the fundamental noise-disturbance connection of quantum measurements takes a very natural form in this framework. Namely, an observable  $\mathbf{A}$  is more noisy than another observable  $\mathbf{B}$  if and only if the set of  $\mathbf{A}$ -channels (the channels that possibly describe the state transformation in some measurement of  $\mathbf{A}$ ) is larger than the set of  $\mathbf{B}$ -channels.

## 8. APPENDIX: ISOMETRIES IN THE PROOF OF THEOREM 1

If  $\dim \mathcal{H} = \infty$ , then the polar decomposition theorem states that a bounded operator  $C$  can be written as  $C = W\sqrt{C^*C}$ , where  $W$  is a partial isometry. Generally,  $W$  cannot be chosen to be isometry. In this Appendix we show that in the two cases treated in Theorem 1, partial isometries can be replaced with isometries.

First, we prove that the operator  $W_x$  in (9) can be chosen to be an isometry. Since  $c_x$  satisfies  $c_x^*c_x = \mathbf{A}(x)$ , there exists a partial isometry  $W_x^0 : \mathcal{H} \rightarrow \mathcal{K}'' \otimes \mathcal{K}'$  satisfying  $c_x = W_x^0\sqrt{\mathbf{A}(x)}$  and  $\text{Ker}[W_x^0] = \text{Ker}[\mathbf{A}(x)]$ . This latter condition implies  $W_x^{0*}W_x^0 = P_{\text{Ker}[\mathbf{A}(x)]^\perp}$  holds, where  $P_{\mathcal{V}}$  for a subspace  $\mathcal{V} \subseteq \mathcal{H}$  is the projection onto  $\mathcal{V}$  and  $\mathcal{V}^\perp$  represents the orthogonal complement of  $\mathcal{V}$ . Let us extend  $W_x^0$  to an

isometry. We have  $\mathbb{1} - A(x) = V^*(\mathbb{1}_{\mathcal{K}'} \otimes (\mathbb{1}_{\mathcal{K}''} - R(x)))V_1$ . Thus there exists a uniquely determined partial isometry  $W'_x$  satisfying

$$(\mathbb{1}_{\mathcal{K}''} \otimes (\mathbb{1}_{\mathcal{K}_1} - R(x))^{1/2})V_1 = W'_x \sqrt{\mathbb{1}_{\mathcal{H}} - A(x)}$$

and  $\text{Ker}[W'_x] = \text{Ker}[\mathbb{1}_{\mathcal{H}} - A(x)]$ . Note that  $\text{Ker}[\mathbb{1}_{\mathcal{H}} - A(x)]^\perp \supseteq \text{Ker}[A(x)]$ . Thus we can restrict  $W'_x$  to  $\text{Ker}[A(x)]$  and write it as  $W_x^1$ . It satisfies  $W_x^{1*}W_x^1 = P_{\text{Ker}[A(x)]}$ . Now it can be shown that  $W^{0*}W^1 = O$ . In fact, we have

$$\begin{aligned} c_x^* d_x P_{\text{Ker}[A(x)]} &= \sqrt{A(x)} W_x^{0*} W_x^1 \sqrt{\mathbb{1}_{\mathcal{H}} - A(x)} P_{\text{Ker}[A(x)]} \\ &= \sqrt{A(x)} W_x^{0*} W_x^1. \end{aligned}$$

The left-hand side of this equality can be written as

$$\begin{aligned} &c_x^* d_x P_{\text{Ker}[A(x)]} \\ &= V^*(\mathbb{1}_{\mathcal{K}'} \otimes R(x))^{1/2} (\mathbb{1}_{\mathcal{K}'} - R(x))^{1/2} V P_{\text{Ker}[A(x)]} \\ &= V^*(\mathbb{1}_{\mathcal{K}'} \otimes (\mathbb{1}_{\mathcal{K}'} - R(x))^{1/2} (\mathbb{1}_{\mathcal{K}'} \otimes R(x))^{1/2}) V P_{\text{Ker}[A(x)]}. \end{aligned}$$

As  $(\mathbb{1}_{\mathcal{K}''} \otimes R(x))^{1/2} V P_{\text{Ker}[A(x)]} = O$  holds, we have  $\sqrt{A(x)} W_x^{0*} W_x^1 = O$  and  $W_x^{0*} W_x^1 = O$ . Thus we can define an isometry  $W_x = W_x^0 \oplus W_x^1$  on the whole space  $\mathcal{H}$ . Consequently we have obtained an isometry  $W_x : \mathcal{H} \rightarrow \mathcal{K}' \otimes \mathcal{K}''$  satisfying  $c_x = W_x \sqrt{A(x)}$ .

Second, we show that the operator  $J_x$  in (11) can be chosen to be an isometry. The relationship  $K^* \hat{A}(x) K = A(x)$  implies that there exists a partial isometry  $J_x^0 : \mathcal{H} \rightarrow \mathcal{K}$  satisfying  $\hat{A}(x) K = J_x^0 \sqrt{A(x)}$  and  $\text{Ker}[J_x^0] = \text{Ker}[A(x)]$ . Since

$$K^*(\mathbb{1} - \hat{A}(x))K = \mathbb{1} - A(x) \quad (16)$$

holds, there exists a partial isometry  $J'_x : \mathcal{H} \rightarrow \mathcal{K}$  satisfying

$$\mathbb{1} - \hat{A}(x) = J'_x \sqrt{\mathbb{1} - A(x)} \quad (17)$$

and  $\text{Ker}[J'_x] = \text{Ker}[\mathbb{1} - A(x)]$ . We denote by  $J_x^1$  the restriction of  $J'_x$  to  $\text{Ker}[A(x)]$ . Then  $J_x := J_x^0 \oplus J_x^1$  is an isometry satisfying  $\hat{A}(x) K = J_x \sqrt{A(x)}$ .

#### ACKNOWLEDGMENTS

TH acknowledges the financial support from the Academy of Finland (grant no. 138135). TM thanks JSPS for the financial support (JSPS KAKENHI Grant Numbers 22740078).

## REFERENCES

- [1] M. Ozawa. Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement. *Phys. Rev. A*, 67:042105, 2003.
- [2] K. Banaszek. Information gain versus state disturbance for a single qubit. *Open Sys. Information Dyn.*, 13:1–16, 2006.
- [3] T. Miyadera and H. Imai. Information-disturbance theorem for mutually unbiased observables. *Phys. Rev. A*, 73:042317, 2006.
- [4] L. Maccone. Entropic information-disturbance tradeoff. *Europhys. Lett.*, 77:40002, 2007.
- [5] F. Buscemi, M. Hayashi, and M. Horodecki. Global information balance in quantum measurements. *Phys. Rev. Lett.*, 100:210504, 2008.
- [6] D. Kretschmann, D. Schlingemann, and R.F. Werner. The information-disturbance tradeoff and the continuity of Stinespring’s representation. *IEEE Trans. Inform. Theory*, 54:1708–1717, 2008.
- [7] M. Ozawa. Operations, disturbance, and simultaneous measurability. *Phys. Rev. A*, 63:032109, 2001.
- [8] T. Heinosaari and M.M. Wolf. Nondisturbing quantum measurements. *J. Math. Phys.*, 51:092201, 2010.
- [9] T. Heinosaari and M. Ziman. *The Mathematical Language of Quantum Theory*. Cambridge University Press, Cambridge, 2012.
- [10] H. Martens and W.M. de Muynck. Nonideal quantum measurements. *Found. Phys.*, 20:255–281, 1990.
- [11] P. Busch, M. Grabowski, and P.J. Lahti. *Operational Quantum Physics*. Springer-Verlag, Berlin, 1997. second corrected printing.
- [12] F. Buscemi, G.M. D’Ariano, M. Keyl, P. Perinotti, and R.F. Werner. Clean positive operator valued measures. *J. Math. Phys.*, 46:082109, 2005.
- [13] T. Heinonen. Optimal measurements in quantum mechanics. *Phys. Lett. A*, 346:77–86, 2005.
- [14] E.B. Davies. *Quantum Theory of Open Systems*. Academic Press, London, 1976.
- [15] M. Ozawa. Quantum measuring processes of continuous observables. *J. Math. Phys.*, 25:79–87, 1984.
- [16] T. Heinosaari, T. Miyadera, and D. Reitzner. Strongly incompatible quantum devices. arXiv:1209.1382 [quant-ph], 2012.
- [17] V. Paulsen. *Completely bounded maps and operator algebras*. Cambridge University Press, Cambridge, 2003.
- [18] W. Arveson. Subalgebras of  $C^*$ -algebras. *Acta Math.*, 123:141–224, 1969.
- [19] M. Raginsky. Radon-Nikodym derivatives of quantum operations. *J. Math. Phys.*, 44:5003–5020, 2003.

‡ TURKU CENTRE FOR QUANTUM PHYSICS, DEPARTMENT OF PHYSICS AND ASTRONOMY, UNIVERSITY OF TURKU

*E-mail address:* teiko.heinosaari@utu.fi

‡ DEPARTMENT OF NUCLEAR ENGINEERING, KYOTO UNIVERSITY, 6068501 KYOTO, JAPAN

*E-mail address:* miyadera@nucleng.kyoto-u.ac.jp